

\*\*\*\*\*  
M P R E H  
(TM)  
\*\*\*\*\*

Release 3.1a John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mprch\_pp protein - protein database search, using Smith-Waterman algorithm  
Rt.: Sun Jun 4 15:31:32 2000; MasPar time 27.68 Seconds  
Tabular output not generated. 698,804 Million cell updates/sec

Title: >US-09-471-749-5  
Description: (1-410) from US09471749.pep  
Perfect Score: 3005  
Sequence: 1 MLDQGVSEYLGVTSEFRKRYP.....QRAPPPRKVGRRGKNSKEG 410

Scoring table: PAM 150  
Gap 11

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: plr62  
1:plr1 2:plr2 3:plr3 4:plr4

Statistics: Mean 48.931; Variance 106.509; scale 0.459

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

R	C	Score	Query Match	Length	ID	Description	Pred. No.
1	436	14.5	371	2	A55302	probable transcriptio	2.00e-53
2	341	11.3	397	2	S26731	neuro-D4 protein - ra	9.49e-37
3	257	8.6	5262	2	T03454	ALR protein - human	1.33e-22
4	253	8.4	4957	2	T03455	ALR protein - human	5.35e-22
5	184	6.1	811	2	T08738	hypothetical protein	4.05e-11
6	176	5.9	1722	2	T78879	retinodastoma bindin	6.25e-10
7	162	5.4	564	2	I48776	Smcy - mouse (fingap	6.75e-08
8	162	5.4	1033	2	I48775	Smcx protein (escap	6.75e-08
9	162	5.4	1560	2	I54361	escapex X-chromosome	6.75e-08
10	160	5.3	728	2	SS7142	hypothetical protein	1.30e-07
11	154	5.1	449	2	T12495	hypothetical protein	9.15e-07
12	149	5.0	1250	2	T00454	hypothetical protein	4.54e-06
13	148	4.9	1214	2	JC2069	zinc-finger protein,	6.24e-06
14	137	4.6	366	2	I53035	trichorax homolog HTX	1.93e-04
15	137	4.6	3869	2	B48205	trichorax homolog HTX	1.93e-04
16	137	4.6	3869	2	A48205	All-1 protein +GRE fo	1.93e-04
17	138	4.6	3910	2	A44264	trichorax homolog HTX	1.42e-04
18	138	4.6	3968	2	A44265	trichorax homolog HTX	1.42e-04
19	134	4.5	288	2	I38485	BCR11Q23 - human (f	4.81e-04
20	133	4.4	653	2	T01274	hypothetical protein	6.50e-04
21	131	4.4	748	2	SS4505	hypothetical protein	1.19e-03
22	133	4.4	825	2	T02518	hypothetical protein	6.50e-04
23	130	4.3	530	2	I38558	Mt-2 autoantigen 240	1.60e-03

24	123	4.1	401	2	S28653	hypothetical protein	1.25e-02
25	123	4.1	1051	2	S55259	trifl protein - mouse	1.25e-02
26	121	4.0	283	2	T11644	hypothetical protein	2.23e-02
27	120	4.0	357	2	JC4090	FK506-binding 39k pro	2.96e-02
28	116	3.9	417	2	S35784	glycoprotein gd - bov	9.17e-02
29	117	3.9	429	2	A40452	keratin 21, type I, c	6.93e-02
30	114	3.8	330	2	S58255	hypothetical protein	1.60e-01
31	112	3.7	933	2	E64486	ATP-dependent 26S pro	4.77e-01
32	110	3.7	933	2	S41539	fibrinogen-binding pr	8.16e-01
33	108	3.6	207	2	T08109	oleosin-like protein	6.25e-01
34	109	3.6	361	2	S61312	cap protein 1, macro	8.16e-01
35	108	3.6	417	1	VGB81B	glycoprotein D precu	1.07e+00
36	107	3.6	852	2	T06310	hypothetical protein	1.07e+00
37	109	3.6	1021	2	S26985	probable DNA-directed	1.07e+00
38	107	3.6	1618	2	S21424	nestin - human	8.16e-01
39	108	3.6	2588	2	T14342	NSD1 protein - mouse	1.39e+00
40	106	3.5	570	2	S52765	secd protein - Strept	1.39e+00
41	106	3.5	596	2	I38228	Shb protein - human	1.39e+00
42	106	3.5	639	2	S39242	Hmg protein - Madagas	1.80e+00
43	105	3.5	651	2	S18874	nucleolin - African c	1.80e+00
44	105	3.5	1191	2	T13850	gene u-shaped protein	1.80e+00
45	105	3.5	1257	2	T01020	hypothetical protein	1.80e+00

## ALIGNMENTS

RESULT 1  
ENTRY A55302 #type complete  
TITLE probable transcription factor regulam - mouse  
ORGANISM Mus musculus #common\_name mouse  
DATE 08-Jul-1995 #sequence\_revision 03-Aug-1995 #text\_change 10-Sep-1997

ACCESSIONS A55302  
REFERENCE A55302  
#authors Gablig, T.G.; Mantel, P.L.; Rosli, R.; Crean, C.D.  
#journal J. Biol. Chem. (1994), 269:29515-29519  
#title Regulam: a novel zinc finger gene essential for apoptosis in myeloid cells.

#cross-references MIMID:95050794  
#accession A55302  
#status preliminary  
#molecule\_type mRNA  
##residues 1-371 #label GAB  
##cross-references GB:010435; NID:9606660; PID:9606661

SUMMARY  
Query Match 14.5%; Score 436; DB 2; Length 371;  
Best Local Similarity 40.3%; Pred. No. 2.00e-53;  
Matches 56; Conservative 31; Mismatches 48; Indels 4; Gaps 3;

DB	216	EEGEDKEDSRPTPVVSORSEQSKKGPDLALPNNYCDFGADSKINKTGQPELVSC	275
QY	256	QSKSKAAK-PRKDGPRKRVLSKVPKRYIPNATGICL-KKESKKRAASLHIC	313
DB	276	SGCGNSGHESCQFTPVMAAVKTYRMWCIIECKCNLCGTSENDQLFCDDCDRGYHMY	335
QY	314	SGCGNSGHESCDLMNEMELYSMTIKTYPMQMECKTCICQAPHEEDMRCDCDGYHMF	373
DB	336	CLTPMSPEPPESNCHLC	354
QY	374	CV-GIGALPSGRWICDC	390

RESULT 2  
ENTRY S26731 #type complete  
TITLE neuro-D4 protein - rat  
ORGANISM Rattus norvegicus #common\_name Norway rat  
DATE 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 17-Mar-1999

ACCESSIONS S26731  
REFERENCE S26731  
#authors Buchman, V.L.; Ninkina, N.N.; Bogdanov, Y.D.; Bortvin, A.L.; Akopian, H.N.; Kiselev, S.L.; Krylova, O.Y.; Anokhin, K.V.;



```

Db      281  PCKMTKMLRN-HSST-QFNMSYVCRCRSDVDKFLKDCGSDNTHIFCLPPLSEV 338
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      322  SCLDWTMEIVSMIKTYPMQCMCKCTCITICGPHHEEMAFCDMCRGYHTFCV-GLGAI 380
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db      339  PKGVWRCPKC 348
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      381  PSGRMICDC 390

RESULT      8
ENTRY      148775
TITLE      Smcx protein (escapes X-chromosome inactivation) - mouse
            (fragment)
ORGANISM    #formal_name Mus musculus domesticus #common_name western
            European house mouse
DATE        12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change
            06-Dec-1996
ACCESSIONS 148775; S44139
REFERENCES 138235
#authors    Agulnik, A.I.; Mitchell, M.J.; Matzel, M.G.; Borsani, G.;
            Ayer, P.A.; Lerner, J.L.; Bishop, C.E.
#journal     Hum. Mol. Genet. (1994) 3:879-884
#title       A novel X gene with a widely transcribed Y-linked homologue
            escapes X-inactivation in mouse and human.
#cross-references MUID:95038739
#accession  148775
#status      Preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues    1-1033 #label RES
#cross-references EMBL:Z29651; NID:g603861; PID:g603862

GENETICS
#gene         Smcx
SUMMARY       #length 1033 #checksum 8222

Query Match
Best Local Similarity 5.4%; Score 162; DB 2; Length 1033;
Matches 24; Conservative 13; Mismatches 29; Indels 4; Gaps 3

Db      265  PCKMTKMLR-RNNHSA-OFISYVCRCRSDDEDKLLKDCGDDNTHIFCLPPLPEI 326
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      323  SCLDWTMEIVSMIKTYPMQCMCKCTCITICGPHHEEMAFCDMCRGYHTFCV-GLGAI 380
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Db      327  PKGVWRCPKC 336
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      381  PSGRMICDC 390

RESULT      9
ENTRY      154361
TITLE      escapes' X-chromosome inactivation - human
            #formal_name Homo sapiens #common_name man
            02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
            28-Feb-1997
ACCESSIONS 154361
REFERENCE 154361
#authors    Wu, J.; Ellison, J.; Salido, E.; Yen, P.; Mohandas, T.;
            Shapiro, L.J.
#journal     Hum. Mol. Genet. (1994) 3:153-160
#title       Isolation and characterization of XE169, a novel human gene
            that escapes X-inactivation.
#cross-references MUID:94214434
#accession  154361
#status      Preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues    1-1560 #label RES
#cross-references GB:L25270; NID:g457136; PID:g457137

GENETICS
#gene         XE169
SUMMARY       #length 1560 #molecular-weight 175804 #checksum 4506

Query Match
Best Local Similarity 5.4%; Score 162; DB 2; Length 1560;

```

```

RESULT 12
ENTRY T00454 #type complete
TITLE hypothetical protein T14N5.11 - Arabidopsis thaliana
ORGANISM #formal_name Arabidopsis thaliana #common_name mouse-ear
cress
DATE 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change
20-Sep-1999
ACCESSIONS T00454
REFERENCE T00454
#authors T14152
#status Federspiel, N.A.; Palm, C.J.; Conway, A.B.; Kurtz, D.B.;
Conway, A.R.; Au, M.; Araujo, R.; Buehler, E.; Dewar, K.;
Feng, J.; Kim, C.; Li, Y.; Ojil, O.; Osborne, B.I.; Shinn,
P.; Sun, H.; Tatum, M.; Vysotskaya, V.S.; Yu, G.; Ecker
J.; Theologis, A.; Davis, R.W.
#submission submitted to the EMBL Data Library, September 1998
#accession T00454
#status preliminary; translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 1-1250 #label FED
#cross-references EMBL:AC004260; NID:93176694; PID:93540206;
GSPDB:GN0059; ATSP:T14N5.11
GENETICS
#gene ATSP:T14N5.11
#map_position 1
#introns 206/3; 256/3; 347/1; 384/1; 437/2; 523/2; 564/2; 590/2;
645/3; 660/3; 692/3; 707/1; 737/3; 790/2; 816/3; 917/3;
947/3
#length 1250 #molecular_weight 141802 #checksum 9025
SUMMARY
Query Match 5.0%; Score 149; DB 2; Length 1250;
Best Local Similarity 29.9%; Pred. No. 4,546-06;
Matches 20; Conservative 16; Mismatches 27; Indels 4; Gaps 4
Db 1126 WYSSC-LCNRCLDCKDDKIVLCGGDAVHYCMRPCEVPNGEPCCTACKAAILIKV 1184
340 WQCECKTCITICGPHHEEMFCDMCRGHTTFCV-GL-GAIFSGRWICDCCORAPPP 397
OY . 398 RKVGRG 404
RESULT 13
ENTRY JC2069 #type complete
TITLE zinc-finger protein, BR140 - human
ALTERNATE_NAMES bromodomain protein
ORGANISM #formal_name Homo sapiens #common_name man
DATE 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change
12-Sep-1997
ACCESSIONS JC2069
REFERENCE JC2069
#authors Thompson, K.A.; Wang, B.; Argaves, W.S.; Glancottl, F.G.;
Schranck, D.P.; Rucslantl, E.
#journal Blochem. Biophys. Res. Commun. (1994) 198:1143-1152
#title BR140, a novel zinc-finger protein with homology to the
TAE250 subunit of TFIIID.
#cross-references MUID:94161726
#accession JC2069
#molecule_type mRNA
#residues 1-1214 #label THO
#cross-references GB:M91585
COMMENT This is a nuclear protein with broad tissue distribution, but is
especially abundant in spermatogonia.
CLASSIFICATION #superfamily unassigned bromodomain proteins; bromodomain
homology
KEYWORDS DNA binding; phosphoprotein; transcription regulation; zinc
finger
FEATURE
653-708 #domain bromodomain homology #label BRO\
23,28,41,44 #binding_site zinc (Cys, Cys, His, His) #status
predicted\

```

```

120,205,462      #binding site phosphate (Ser)(covalent) (by casein
276,279,293,296 kinase II) #status predicted\
301,304,317,320 #binding_site zinc (Cys) #status predicted\
330,333,350,353 #binding_site zinc (His, Cys, Cys) #status
386,389,401,405 #binding_site zinc (Cys, Cys, His, Cys) #status
410,413,444,447 #binding_site zinc (His, Cys, Cys, His) #status
predicted
#length 1214 #molecular-weight 137484 #checksum 1329
SUMMARY
Query Match 4.9%; Score 148; DB 2; Length 1214;
Best Local Similarity 38.0%; Pred. No. 6,246-06;
Matches 19; Conservative 14; Mismatches 15; Indels 2; Gaps 1;
D 276 CCIQNDGCONSNVILFCMDNCLVHQCVCVPYIPSGOMICRCRLQSPS 325
C 348 CIIIC-GOPHHEEMKFCMDCRGRTHTFCVGLGAIIPSGRWICDCCGRAP 395
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
RESULT 14
ENTRY 153035 #type fragment
TITLE trichorax homolog HIX - human (fragment)
ORGANISM #format_name Homo sapiens #common_name man
DATE 01-Nov-1996 #sequence_revision 01-Nov-1996 #text_change
29-Aug-1997
ACCESSIONS 153035
REFERENCE 153035
AUTHORS Mbangkollo, D.; Burnett, R.; McCabe, N.; Thirman, M.; Gill,
H.; Yu, H.; Rowley, J.D.; Diaz, M.O.
#journal DNA Cell Biol. (1995) 14:475-483
#title The human MTL gene: nucleotide sequence, homology to the
Drosophila trx zinc-finger domain, and alternative
splicing
#cross-references M01D:95322025
#accession I53035
#status Preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-366 #label RES
#cross-references GB:S78570; NID:g1042096; PID:g1042097
GENETICS
#gene GDB:MLT
#cross-references GDB:128819; OMIM:159555
#map_position 11q23-11q23
SYNCRY #length 366 #checksum 7616
#entry Match 4.6%; Score 137; DB 2; Length 366;
Best Local Similarity 29.0%; Pred. No. 1,936-04;
Matches 20; Conservative 16; Mismatches 29; Indels 4; Gaps 3
D 157 FYVQVOCPEPHKPLENERPLEDQLEN--WCGRRCFCYVCGRQNRHATQLLECKNCR 214
OY 310 LIHCQCSNSHSPCLDWT-MELVSMKRTYVQMCCKCTIICGQPHHE-BEMKFCMDCD 367
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
DB 215 NSYAPFCLG 223
OY 368 RGYHFCVYG 376
| | | | | : : : : : | | | | | : : : : : | | | | | : : : : :
RESULT 15
ENTRY B48205 #type fragment
TITLE All-1 protein -GFE form - mouse (fragment)
ORGANISM #format_name Mus musculus #common_name house mouse
DATE 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change
07-Oct-1994
ACCESSIONS B48205
REFERENCE A48205
AUTHORS Ma, O.; Alder, H.; Nelson, K.R.; Chatterjee, D.; Gu, Y.;
Nakamura, T.; Canaan, E.; Croce, C.M.; Stracusa, L.D.;
Buchberg, A.M.
#journal Proc. Natl. Acad. Sci. U.S.A. (1993) 90:6350-6354

```

```

#title      Analysis of the murine All-1 gene reveals conserved domains
            with human All-1 and identifies a motif shared with DNA
            methyltransferases.
#cross-references MUID:93317679
#accession  B48205
#status     preliminary
#molecule_type mRNA
#residues   1-3866 #label HAR
##cross-references GB:L17069

GENETICS
#gene       All-1
KEYWORDS    alternative splicing; zinc finger
SUMMARY     #length 3866 #checksum 2325

Query Match           4.6%; Score 137; DB 2; Length 3866;
Best Local Similarity 29.0%; Pred. No. 1.93e-04;
Matches 20; Conservative 16; Mismatches 29; Indels 4; Gaps 3

Db 1347 FVYQVCCEPRHKKCLENERRPDLQLEN--WCCRRCFCFVCGRQATQTQLLECKKR 1404
        ::::: |::: : | | | |::: |::: |:::
Qy 310 LIHSCQCNSSGHSPCLDWT-MELVSMIKTYPMQCMEKTCIIIGOPNHE-BEMAFCDMCD 367
        ::::: ||| |::: |::: |::: |::: |:::
Db 1405 NSYHPCELG 1413
Qy 368 RGYHTFCVG 376

```

\*\*\*\*\*  
M P E R E H  
(TM)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Msrch\_pp protein - protein database search, using Smith-Waterman algorithm  
R: Sun Jun 4 15:32:58 2000; Maspar time 11.26 Seconds  
525.809 Million cell updates/sec  
Tabular output not generated.

Title: >US-09-471-749-5  
Description: (1-410) from US09471749.pep  
Perfect Score: 3005  
Sequence: 1 MDAQVSEYLGVTSEFRKYP.....QRPAPPRVGRGRKSKG 410

Scoring table:  
PAM 150  
Gap 11

Searched: 145341 segs, 14437480 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-issued  
1:5L\_COMB 2:5B\_COMB 3:6\_COMB 4:PC1\_COMB 5:backfiles1

Statistics: Mean 33.110; Variance 152.561; scale 0.217  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Score	Query	Match	Length	ID	Description	Pred. No.
441	14.7	405	2	US-08-881-	Sequence 2, Applicatio	7.37e-28
147	4.9	113	4	PCR-US94-0	Sequence 37, Applicati	4.88e-03
147	4.9	113	1	US-08-320-	Sequence 37, Applicati	4.88e-03
147	4.9	113	1	US-08-545-	Sequence 37, Applicati	4.88e-03
146	4.9	187	3	US-08-545-	Sequence 59, Applicati	5.82e-03
146	4.9	187	4	PCR-US94-0	Sequence 59, Applicati	5.82e-03
137	4.6	1400	4	PCR-US93-0	Sequence 59, Applicati	5.82e-03
137	4.6	1400	1	US-08-080-	Sequence 7, Applicatio	2.81e-02
130	4.3	351	4	PCR-US91-0	Sequence 1, Applicatio	9.39e-02
118	3.9	373	2	US-08-864-	Sequence 3, Applicatio	7.16e-01
110	3.7	933	3	US-08-293-	Sequence 2, Applicatio	2.69e+00
108	3.6	417	2	US-08-682-	Sequence 6, Applicatio	3.72e+00
107	3.6	1618	1	PCR-US83-	Sequence 4, Applicatio	4.38e+00
104	3.5	218	4	PCR-US94-0	Sequence 2, Applicatio	7.11e+00
104	3.5	218	1	US-08-327-	Sequence 2, Applicatio	7.11e+00
104	3.5	218	1	US-08-320-	Sequence 2, Applicatio	7.11e+00
104	3.5	218	1	US-08-545-	Sequence 2, Applicatio	6.05e+00
105	3.5	1376	1	PCR-US93-1	Sequence 3, Applicatio	6.05e+00
105	3.5	1376	4	PCR-US93-1	Sequence 3, Applicatio	6.05e+00
103	3.4	188	4	PCR-US94-0	Sequence 58, Applicati	8.35e+00
103	3.4	188	4	PCR-US94-0	Sequence 55, Applicati	8.35e+00
103	3.4	1093	4	PCR-US94-0	Sequence 55, Applicati	8.35e+00
103	3.4	1093	3	US-08-545-	Sequence 55, Applicati	8.35e+00

RESULT	ID	US-08-881-857-2	STANDARD;	PRT;	405 AA.
24	98	3.3	781	2	US-08-675- Sequence 3, Applicatio 1.85e+01
25	98	3.3	788	2	US-08-907- Sequence 6, Applicatio 1.85e+01
26	98	3.3	997	3	US-08-872- Sequence 8, Applicatio 1.85e+01
27	96	3.2	210	2	US-09-258- Sequence 2, Applicatio 2.53e+01
28	96	3.2	210	3	US-08-569- Sequence 2, Applicatio 2.53e+01
29	96	3.2	210	2	US-09-258- Sequence 2, Applicatio 2.53e+01
30	96	3.2	250	2	US-08-867- Sequence 13, Applicati 2.96e+01
31	96	3.2	294	2	US-09-258- Sequence 10, Applicati 2.53e+01
32	96	3.2	414	1	US-07-667- Sequence 4, Applicatio 2.53e+01
33	96	3.2	593	1	US-08-296- Sequence 2, Applicatio 2.53e+01
34	96	3.2	673	2	US-08-455- Sequence 6, Applicatio 2.53e+01
35	96	3.2	704	1	US-08-188- Sequence 18, Applicati 2.53e+01
36	96	3.2	704	1	US-08-646- Sequence 18, Applicati 2.53e+01
37	96	3.2	806	1	US-07-980- Sequence 2, Applicatio 2.96e+01
38	95	3.2	861	2	US-08-343- Sequence 18, Applicati 2.96e+01
39	95	3.2	1958	1	US-07-945- Sequence 2, Applicatio 4.70e+01
40	93	3.1	104	2	US-08-710- Sequence 19, Applicati 4.70e+01
41	92	3.1	182	2	US-08-466- Sequence 90, Applicati 4.70e+01
42	92	3.1	546	2	US-09-067- Sequence 1, Applicatio 4.70e+01
43	92	3.1	651	3	US-08-650- Sequence 6, Applicatio 4.70e+01
44	92	3.1	1070	3	US-08-922- Sequence 22, Applicati 4.70e+01
45	92	3.1	1070	2	US-08-633- Sequence 2, Applicatio 4.70e+01

ALIGNMENTS

Sequence 1  
US-08-881-857-2  
xxxxxx

Sequence 2, Application US/08881857

Patent No. 5919660  
GENERAL INFORMATION:  
APPLICANT: KIKLY, KRISTINE R  
APPLICANT: GROSS, MITCHELL S  
APPLICANT: HURLE, MARK R  
TITLE OF INVENTION: HUMAN REQULEM  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: RATNER & PRESTIA  
STREET: P.O. BOX 980  
CITY: VALLEY Forge  
STATE: PA  
COUNTRY: USA  
ZIP: 19482

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/881,857  
FILING DATE: 24-JUN-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/021,229  
FILING DATE: 26-JUN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: PRESTIA, PAUL F  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: ATG-50013  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0701  
TELEX: 846169  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 405 amino acids



Query Match 4.9%; Score 147; DB 1; Length 113;  
Best Local Similarity 29.1%; Pred. No. 4,88e-03;  
Matches 23; Conservative 19; Mismatches 32; Indels 5; Gaps 4;

Db 17 ENISKSD-EVYCCOVCEPFHFKLENERPLEDQLEN--WCRCRKFCHVCGRHOAT 73  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Cc 300 ESNKKGAESLIHCSCNSGHPSCLDWT-MEIVSMITTYMOCKECTCIIICGPHE- 357

Db 74 KOLECNKCNRNSYHEPCIG 92  
:: | :: | : | : | : | : | : | : | : | : | : | : | : | : | :  
Cc 358 EMMFCDMDCRGYHTFCVG 376

RESULT 4 STANDARD; PRT; 113 AA.  
ID US-08-545-86OD-37  
AC xxxxxx  
XX Sequence 37, Application US/0854586OD  
DE Patent No. 6040140  
CC GENERAL INFORMATION:  
CC APPLICANT: Croce, Carlo  
CC APPLICANT: Canaanl, Eli  
CC TITLE OF INVENTION: Diagnostics, Therapeutics and Methods  
CC TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias  
CC TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the ALL-1 Regio  
CC NUMBER OF SEQUENCES: 94  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Woodcock, Washburn, Kurtz, Macklewitz &  
CC ADDRESSEE: No. 604014Oris  
CC STREET: One Liberty Place, 46th floor  
CC CITY: Philadelphia  
CC STATE: Pennsylvania  
CC COUNTRY: USA  
CC ZIP: 19103  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
Cc APPLICATION NUMBER: US/08/545,86OD  
Cc FILING DATE: 07-MAR-1996  
Cc CLASSIFICATION: 435  
Cc PRIOR APPLICATION DATA:  
Cc APPLICATION NUMBER: PCT/US94/04496  
Cc FILING DATE: 22-APR-1994  
Cc PRIOR APPLICATION DATA:  
Cc APPLICATION NUMBER: PCT/US92/10930  
Cc FILING DATE: 09-DEC-1992  
Cc PRIOR APPLICATION DATA:  
Cc APPLICATION NUMBER: US 08/327,392  
Cc FILING DATE: 19-OCT-1994  
Cc PRIOR APPLICATION DATA:  
Cc APPLICATION NUMBER: US 08/320,559  
Cc FILING DATE: 11-OCT-1994  
Cc PRIOR APPLICATION DATA:  
Cc APPLICATION NUMBER: US 08/062,443  
Cc FILING DATE: 14-MAY-1993  
Cc PRIOR APPLICATION DATA:  
Cc APPLICATION NUMBER: US 07/971,094  
Cc FILING DATE: 30-OCT-1992  
Cc PRIOR APPLICATION DATA:  
Cc APPLICATION NUMBER: US 07/888,839  
Cc FILING DATE: 27-MAY-1992  
Cc PRIOR APPLICATION DATA:  
Cc APPLICATION NUMBER: US 07/805,093  
Cc FILING DATE: 11-DEC-1991

```
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Deluca Esq., Mark  
CC REGISTRATION NUMBER: 33,229  
CC REFERENCE/DOCKET NUMBER: TJU-1262  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (215) 568-3100  
CC TELEFAX: (215) 568-3439  
CC INFORMATION FOR SEQ ID NO: 37:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 113 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 113 AA; 13317 MW; 64891 CN;  
  
Query Match 4.9%; Score 147; DB 3; Length 113;  
Best Local Similarity 29.1%; Pred. No. 4,88e-03;  
Matches 23; Conservative 19; Mismatches 32; Indels 5; Gaps 4;  
  
Db 17 EENISSKSD-FYVQVCCEPHRFCLSENERPLEDOLFN--WCCRCKFCFHCGRHOAT 73  
QY 300 ESNKKAESLIHSCGENSEGHPSCLDMT-MELYSMIKTYPWOCMECKTCIIICGPARE- 357  
:::|::|::|::|::|  
Db 74 KOLECNCRCRSHPECLG 92  
QY 358 EEMMFCDMDCRGHFTCVG 376  
:::|::|::|::|::|  
  
RESULT 5 STANDARD; PRT; 187 AA.  
ID US-08-545-860D-59  
AC xxxxxx  
DT  
XX  
DE Sequence 59, Application US/08545860D  
CC  
CC Patent No. 6040140  
CC GENERAL INFORMATION:  
CC APPLICANT: Croce, Carlo  
CC APPLICANT: Canaan, Eli  
CC TITLE OF INVENTION: Diagnostics, Therapeutics and Methods  
CC TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias  
CC TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1 Re  
CC NUMBER OF SEQUENCES: 94  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: Woodcock, Washburn, Kurtz, Macklewitz &  
CC ADDRESSEE: No. 604014Oris  
CC STREET: One Liberty Place, 46th floor  
CC City: Philadelphia  
CC STATE: Pennsylvania  
CC COUNTRY: USA  
CC ZIP: 19103  
CC  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/545,860D  
CC FILING DATE: 07-MAR-1996  
CC CLASSIFICATION: 435,  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US94/04496  
CC FILING DATE: 22-APR-1994  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US92/10930  
CC FILING DATE: 09-DEC-1992  
CC PRIOR APPLICATION DATA:  
CC APPLICATION NUMBER: US 08/327,392  
CC FILING DATE: 19-OCT-1994  
CC PRIOR APPLICATION DATA:
```



CC		APPLICATION NUMBER:	US 08/320,559
CC		FILING DATE:	11-OCT-1994
CC		PRIOR APPLICATION DATA:	
CC		APPLICATION NUMBER:	US 08/062,443
CC		FILING DATE:	14-MAY-1993
CC		PRIOR APPLICATION DATA:	
CC		APPLICATION NUMBER:	US 07/971,094
CC		FILING DATE:	30-OCT-1992
CC		PRIOR APPLICATION DATA:	
CC		APPLICATION NUMBER:	US 07/888,839
CC		FILING DATE:	27-MAY-1992
CC		PRIOR APPLICATION DATA:	
CC		APPLICATION NUMBER:	US 07/805,093
CC		FILING DATE:	11-DEC-1991
CC		ATTORNEY/AGENT INFORMATION:	
CC		NAME:	Deluca Esq., Mark
CC		REGISTRATION NUMBER:	33,229
CC		REFERENCE/DOCKET NUMBER:	TJ0-1262
CC		TELECOMMUNICATION INFORMATION:	
CC		TELEPHONE:	(215) 568-3100
CC		TELEFAX:	(215) 568-3439
CC		INFORMATION FOR SEQ ID NO:	59:
CC		SEQUENCE CHARACTERISTICS:	
CC		LENGTH:	187 amino acids
CC		TYPE:	amino acid
CC		STRANDEDNESS:	single
CC		TOPOLOGY:	linear
CC		MOLECULE TYPE:	protein
CC		HYPOTHETICAL:	NO
CC		ANTI-SENSE:	NO
SQ		SEQUENCE	187 AA; 20805 MW; 171203 CN;
Df		Query Match	4.9%; Score 146; DB 3; Length 187;
Df		Best Local Similarity	38.0%; Pred.No.5,82e-03;
Matches	19;	Conservative	13; Mismatches 16; Indels 2; Gaps 1;
Df	8	CCICNDGCONSNVILPCDMCNLEVDCEYVPPIPSQMLCRCLOSPS	57
Qy	348	CITC--GGPHHEEMFCMCDRGVHTFCVGIGAIPSGRMWICDCGRAP	395
RESULT	6		
ID	PCT-US94-04496-59	STANDARD;	PRT: 187 AA.
AC	xxxxxx		
DT			
XX			
DF	Sequence 59, Application PC/TUS9404496		
CC	GENERAL INFORMATION:		
CC	APPLICANT:	Croce, Carlo	
CC	APPLICANT:	Canaan, Eli	
CC	TITLE OF INVENTION:	Diagnostics, Therapeutics and Methods	
CC	TITLE OF INVENTION:	for Detection and Treatment of Acute Leukemias	
CC	TITLE OF INVENTION:	Resulting from Chromosome Abnormalities in the All-I-	
CC	NUMBER OF SEQUENCES:	86	
CC	CORRESPONDENCE ADDRESS:		
CC	ADDRESSEE:	Woodcock, Washburn, Kurtz, Macklewitz &	
CC	ADDRESSER:	Norris	
CC	STREET:	One Liberty Place, 46th floor	
CC	CITY:	Philadelphia	
CC	STATE:	Pennsylvania	
CC	COUNTRY:	USA	
CC	ZIP:	19103	
CC	COMPUTER READABLE FORM:		
CC	MEDIUM TYPE:	Floppy disk	
CC	COMPUTER:	IBM PC compatible	
CC	OPERATING SYSTEM:	PC-DOS/MS-DOS	
CC	SOFTWARE:	PatentIn Release #1.0, Version #1.25	
CC	CURRENT APPLICATION DATA:		
CC	APPLICATION NUMBER:	PCT/US94/04496	

```
CC      FILING DATE:  
CC      CLASSIFICATION:  
CC      ATTORNEY/AGENT INFORMATION:  
CC          NAME: Deluca Esq., Marx  
CC          REGISTRATION NUMBER: 33, 229  
CC          REFERENCE/DOCKET NUMBER: TJU-1242  
CC          TELECOMMUNICATION INFORMATION:  
CC              TELEPHONE: (215) 568-3100  
CC              TELEFAX: (215) 568-3439  
CC          INFORMATION FOR SEQ ID NO: 59:  
CC              SEQUENCE CHARACTERISTICS:  
CC                  LENGTH: 187 amino acids  
CC                  TYPE: amino acid  
CC                  STRANDEDNESS: single  
CC                  TOPOLOGY: linear  
CC          MOLECULE TYPE: protein  
CC          HYPOCHETICAL: NO  
CC          ANTI-SENSE: NO  
SQ      SEQUENCE 187 AA; 20805 MW; 171203 CN;  
  
Cc      Query Match                4.9%; Score 146; DB 4; Length 187;  
Cg      Best Local Similarity    38.0%; Pred.No. 5.82e-03;  
Matches   19; Conservative     13; Mismatches 16; Indels 2; Caps 1;  
  
Db      8 CCICDGCNSNVLFCDCMCNLNEVHQCQGVPIIPGCGMLCRCLOSP 57  
        ||| : ::::||| | : | : | : | : | :  
Qy      348 CII-C-GQPHEEMMFCMDRCRGYHTFGVGGLAIFSGRMICCDGORAP 395  
  
RESULT       7  
ID            PCT-US93--05857-7           STANDARD;         PRI;      1400 AA.  
xx             xxxxxx  
AC  
AT  
DT  
XX  
Sequence 7, Application PC/TUS9305857  
DE  
xx  
Sequence 7, Application PC/TUS9305857  
cc  
GENERAL INFORMATION:  
cc  
APPLICANT: Board of Regents  
cc  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DETECTING  
cc  
GENE REARRANGEMENTS AND TRANSLOCATIONS  
cc  
NUMBER OF SEQUENCES: 8  
cc  
CORRESPONDENCE ADDRESS:  
cc  
ADDRESSEE: Arnold, White & Durkee  
cc  
STREET: P. O. Box 4433  
cc  
CITY: Houston  
cc  
STATE: Texas  
cc  
COUNTRY: USA  
cc  
ZIP: 77210  
cc  
COMPUTER READABLE FORM:  
cc  
MEDIUM TYPE: Floppy disk  
cc  
COMPUTER: IBM PC compatible  
cc  
OPERATING SYSTEM: PC-DOS/MS-DOS  
cc  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
cc  
CURRENT APPLICATION DATA:  
cc  
APPLICATION NUMBER: PCT/US93/05857  
cc  
FILING DATE: 19930617  
cc  
CLASSIFICATION:  
cc  
PRIOR APPLICATION DATA:  
cc  
APPLICATION NUMBER: 07/900,689  
cc  
FILING DATE: 17/06/92  
cc  
ATTORNEY/AGENT INFORMATION:  
cc  
NAME: Parker, David L.  
cc  
REGISTRATION NUMBER: 32,165  
cc  
REFERENCE/DOCKET NUMBER: ARCD:072/PAR  
cc  
TELECOMMUNICATION INFORMATION:  
cc  
TELEPHONE: (512) 320-7200  
cc  
TELEFAX: (512) 474-7577  
cc  
INFORMATION FOR SEQ ID NO: 7:  
cc  
SEQUENCE CHARACTERISTICS:
```









QY 116 WOPSEÖGTRHPL 128

XXXXXX

Sequence 32, Application, US/08680326

CORRESPONDENCE ADDRESS:

STREET: 755 Page Mill Road

STATE: California

ZIP: 94304-1018

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

0-14 DECEMBER 1961

RENT APPLICATION DATA:

FILE DATE: 02/06/00

ORNEY/AGENT INFORMATION:

REGISTRATION NUMBER: 40,253

TELEPHONE: 74151 813 5600

ELFAX: (415) 494-0792  
ELFAX: 706141

ATION FOR SEQ ID NO: 32:

LENGTH: 1015 amino acids  
TYPE: amino acid

TRANDEDNESS: single

1015 AA; 113417 MW; 5145857 C

Similarity 45.5%: Pred No 3  
Score 87; D

10; Conservative 6; Mismatch

96T 9GWSIMGYG 196

29 JAN 1962

PRT; STANDARD;

\*\*\*\*\*  
M P E R E H  
(7M)  
\*\*\*\*\*

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

MP\_hrp protein - protein database search, using Smith-Waterman algorithm  
Run: Sun Jun 4 15:28:16 2000; MasPar time 7.59 Seconds  
452.598 Million cell updates/sec  
Tabular output not generated.

Title: >US-09-471-749-3  
Description: (1-238) from US09471749.pep  
Perfect Score: 1742  
Sequence: 1 MEYMKRSLSLYPRSLSRHV.....GPPKGRKASLIPCTIKTIQ 238

Scoring table: PAM 150  
Gap 11

Searched: 145341 segs, 14437480 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: a-issued  
1:5A\_COMB 2:5B\_COMB 3:6\_COMB 4:PCT\_COMB 5:backfilest1

Statistics: Mean 30.727; Variance 132.659; scale 0.232

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Re. No.	Score	Query Match	Length	ID	Description	Pred. No.
1	95	5.5	403	4	PCT-US95-0	Sequence 2, Applicatio 9.91e+00
2	95	5.5	403	3	US-08-989-	Sequence 4, Applicatio 9.91e+00
3	95	5.5	403	2	US-08-634-	Sequence 4, Applicatio 9.91e+00
4	95	5.5	403	2	US-08-634-	Sequence 2, Applicatio 9.91e+00
5	95	5.5	403	2	US-08-607-	Sequence 4, Applicatio 9.91e+00
6	95	5.5	403	2	US-08-607-	Sequence 2, Applicatio 9.91e+00
7	95	5.5	403	2	US-08-533-	Sequence 10, Applicatio 9.91e+00
8	95	5.5	403	2	US-08-454-	Sequence 2, Applicatio 9.91e+00
9	95	5.5	403	3	US-08-989-	Sequence 5, Applicatio 2.73e+01
10	89	5.1	431	2	US-08-665-	Sequence 32, Applicatio 3.81e+01
11	87	5.0	1015	2	US-08-680-	Sequence 6, Applicatio 4.49e+01
12	87	5.0	1130	2	US-08-519-	Sequence 89, Applicatio 4.49e+01
13	86	4.9	376	2	US-08-997-	Sequence 89, Applicatio 4.49e+01
14	86	4.9	376	2	US-08-997-	Sequence 89, Applicatio 4.49e+01
15	86	4.9	376	2	US-08-997-	Sequence 89, Applicatio 4.49e+01
16	85	4.9	497	1	US-08-633-	Sequence 6, Applicatio 5.29e+01
17	85	4.9	497	1	US-08-295-	Sequence 6, Applicatio 5.29e+01
18	85	4.9	546	2	US-08-713-	Sequence 16, Applicatio 7.34e+01
19	83	4.8	445	2	PCT-US94-0	Sequence 9, Applicatio 7.34e+01
20	83	4.8	170	4	US-08-470-	Sequence 9, Applicatio 7.34e+01
21	83	4.8	170	1	US-08-081-	Sequence 8, Applicatio 7.34e+01
22	83	4.8	170	1	US-08-081-	Sequence 14, Applicatio 7.34e+01
23	83	4.8	233	3	US-08-471-	Sequence 14, Applicatio 7.34e+01

24	83	4.8	233	1	US-08-081-	Sequence 6, Applicatio 7.34e+01
25	83 <td>4.8<td>233<td>4<td>PCT-US95-0</td><td>Sequence 24, Applicatio 7.34e+01</td></td></td></td>	4.8 <td>233<td>4<td>PCT-US95-0</td><td>Sequence 24, Applicatio 7.34e+01</td></td></td>	233 <td>4<td>PCT-US95-0</td><td>Sequence 24, Applicatio 7.34e+01</td></td>	4 <td>PCT-US95-0</td> <td>Sequence 24, Applicatio 7.34e+01</td>	PCT-US95-0	Sequence 24, Applicatio 7.34e+01
26	83 <td>4.8<td>233<td>2<td>PCT-US94-0</td><td>Sequence 7, Applicatio 7.34e+01</td></td></td></td>	4.8 <td>233<td>2<td>PCT-US94-0</td><td>Sequence 7, Applicatio 7.34e+01</td></td></td>	233 <td>2<td>PCT-US94-0</td><td>Sequence 7, Applicatio 7.34e+01</td></td>	2 <td>PCT-US94-0</td> <td>Sequence 7, Applicatio 7.34e+01</td>	PCT-US94-0	Sequence 7, Applicatio 7.34e+01
27	83 <td>4.8<td>233<td>2<td>US-08-661-</td><td>Sequence 59, Applicatio 7.34e+01</td></td></td></td>	4.8 <td>233<td>2<td>US-08-661-</td><td>Sequence 59, Applicatio 7.34e+01</td></td></td>	233 <td>2<td>US-08-661-</td><td>Sequence 59, Applicatio 7.34e+01</td></td>	2 <td>US-08-661-</td> <td>Sequence 59, Applicatio 7.34e+01</td>	US-08-661-	Sequence 59, Applicatio 7.34e+01
28	83 <td>4.8<td>233<td>1<td>US-08-607-</td><td>Sequence 24, Applicatio 7.34e+01</td></td></td></td>	4.8 <td>233<td>1<td>US-08-607-</td><td>Sequence 24, Applicatio 7.34e+01</td></td></td>	233 <td>1<td>US-08-607-</td><td>Sequence 24, Applicatio 7.34e+01</td></td>	1 <td>US-08-607-</td> <td>Sequence 24, Applicatio 7.34e+01</td>	US-08-607-	Sequence 24, Applicatio 7.34e+01
29	83 <td>4.8<td>233<td>1<td>US-08-471-</td><td>Sequence 14, Applicatio 7.34e+01</td></td></td></td>	4.8 <td>233<td>1<td>US-08-471-</td><td>Sequence 14, Applicatio 7.34e+01</td></td></td>	233 <td>1<td>US-08-471-</td><td>Sequence 14, Applicatio 7.34e+01</td></td>	1 <td>US-08-471-</td> <td>Sequence 14, Applicatio 7.34e+01</td>	US-08-471-	Sequence 14, Applicatio 7.34e+01
30	83 <td>4.8<td>233<td>2<td>US-08-470-</td><td>Sequence 7, Applicatio 7.34e+01</td></td></td></td>	4.8 <td>233<td>2<td>US-08-470-</td><td>Sequence 7, Applicatio 7.34e+01</td></td></td>	233 <td>2<td>US-08-470-</td><td>Sequence 7, Applicatio 7.34e+01</td></td>	2 <td>US-08-470-</td> <td>Sequence 7, Applicatio 7.34e+01</td>	US-08-470-	Sequence 7, Applicatio 7.34e+01
31	83 <td>4.8<td>233<td>1<td>US-08-333-</td><td>Sequence 59, Applicatio 7.34e+01</td></td></td></td>	4.8 <td>233<td>1<td>US-08-333-</td><td>Sequence 59, Applicatio 7.34e+01</td></td></td>	233 <td>1<td>US-08-333-</td><td>Sequence 59, Applicatio 7.34e+01</td></td>	1 <td>US-08-333-</td> <td>Sequence 59, Applicatio 7.34e+01</td>	US-08-333-	Sequence 59, Applicatio 7.34e+01
32	83 <td>4.8<td>258<td>1<td>US-07-990-</td><td>Sequence 2, Applicatio 6.23e+01</td></td></td></td>	4.8 <td>258<td>1<td>US-07-990-</td><td>Sequence 2, Applicatio 6.23e+01</td></td></td>	258 <td>1<td>US-07-990-</td><td>Sequence 2, Applicatio 6.23e+01</td></td>	1 <td>US-07-990-</td> <td>Sequence 2, Applicatio 6.23e+01</td>	US-07-990-	Sequence 2, Applicatio 6.23e+01
33	84 <td>4.8<td>371</td><td>2<td>US-08-378-</td><td>Sequence 8, Applicatio 6.23e+01</td></td></td>	4.8 <td>371</td> <td>2<td>US-08-378-</td><td>Sequence 8, Applicatio 6.23e+01</td></td>	371	2 <td>US-08-378-</td> <td>Sequence 8, Applicatio 6.23e+01</td>	US-08-378-	Sequence 8, Applicatio 6.23e+01
34	84 <td>4.8<td>749</td><td>3<td>US-08-956-</td><td>Sequence 8, Applicatio 6.23e+01</td></td></td>	4.8 <td>749</td> <td>3<td>US-08-956-</td><td>Sequence 8, Applicatio 6.23e+01</td></td>	749	3 <td>US-08-956-</td> <td>Sequence 8, Applicatio 6.23e+01</td>	US-08-956-	Sequence 8, Applicatio 6.23e+01
35	84 <td>4.8<td>749</td><td>4<td>PCT-US95-1</td><td>Sequence 8, Applicatio 6.23e+01</td></td></td>	4.8 <td>749</td> <td>4<td>PCT-US95-1</td><td>Sequence 8, Applicatio 6.23e+01</td></td>	749	4 <td>PCT-US95-1</td> <td>Sequence 8, Applicatio 6.23e+01</td>	PCT-US95-1	Sequence 8, Applicatio 6.23e+01
36	84 <td>4.8<td>749</td><td>2<td>US-08-852-</td><td>Sequence 8, Applicatio 6.23e+01</td></td></td>	4.8 <td>749</td> <td>2<td>US-08-852-</td><td>Sequence 8, Applicatio 6.23e+01</td></td>	749	2 <td>US-08-852-</td> <td>Sequence 8, Applicatio 6.23e+01</td>	US-08-852-	Sequence 8, Applicatio 6.23e+01
37	84 <td>4.8<td>749</td><td>2<td>US-08-820-</td><td>Sequence 8, Applicatio 6.23e+01</td></td></td>	4.8 <td>749</td> <td>2<td>US-08-820-</td><td>Sequence 8, Applicatio 6.23e+01</td></td>	749	2 <td>US-08-820-</td> <td>Sequence 8, Applicatio 6.23e+01</td>	US-08-820-	Sequence 8, Applicatio 6.23e+01
38	84 <td>4.8<td>749</td><td>1<td>US-08-369-</td><td>Sequence 13, Applicatio 6.23e+01</td></td></td>	4.8 <td>749</td> <td>1<td>US-08-369-</td><td>Sequence 13, Applicatio 6.23e+01</td></td>	749	1 <td>US-08-369-</td> <td>Sequence 13, Applicatio 6.23e+01</td>	US-08-369-	Sequence 13, Applicatio 6.23e+01
39	83 <td>4.8<td>1159</td><td>2<td>US-08-470-</td><td>Sequence 20, Applicatio 8.63e+01</td></td></td>	4.8 <td>1159</td> <td>2<td>US-08-470-</td><td>Sequence 20, Applicatio 8.63e+01</td></td>	1159	2 <td>US-08-470-</td> <td>Sequence 20, Applicatio 8.63e+01</td>	US-08-470-	Sequence 20, Applicatio 8.63e+01
40	82 <td>4.7<td>462</td><td>2<td>US-08-471-</td><td>Sequence 20, Applicatio 8.63e+01</td></td></td>	4.7 <td>462</td> <td>2<td>US-08-471-</td><td>Sequence 20, Applicatio 8.63e+01</td></td>	462	2 <td>US-08-471-</td> <td>Sequence 20, Applicatio 8.63e+01</td>	US-08-471-	Sequence 20, Applicatio 8.63e+01
41	82 <td>4.7<td>462</td><td>2<td>US-08-471-</td><td>Sequence 20, Applicatio 8.63e+01</td></td></td>	4.7 <td>462</td> <td>2<td>US-08-471-</td><td>Sequence 20, Applicatio 8.63e+01</td></td>	462	2 <td>US-08-471-</td> <td>Sequence 20, Applicatio 8.63e+01</td>	US-08-471-	Sequence 20, Applicatio 8.63e+01
42	82 <td>4.7<td>462</td><td>2<td>US-08-471-</td><td>Sequence 20, Applicatio 8.63e+01</td></td></td>	4.7 <td>462</td> <td>2<td>US-08-471-</td><td>Sequence 20, Applicatio 8.63e+01</td></td>	462	2 <td>US-08-471-</td> <td>Sequence 20, Applicatio 8.63e+01</td>	US-08-471-	Sequence 20, Applicatio 8.63e+01
43	82 <td>4.7<td>462</td><td>1<td>US-08-463-</td><td>Sequence 20, Applicatio 8.63e+01</td></td></td>	4.7 <td>462</td> <td>1<td>US-08-463-</td><td>Sequence 20, Applicatio 8.63e+01</td></td>	462	1 <td>US-08-463-</td> <td>Sequence 20, Applicatio 8.63e+01</td>	US-08-463-	Sequence 20, Applicatio 8.63e+01
44	82 <td>4.7<td>462</td><td>1<td>US-08-471-</td><td>Sequence 20, Applicatio 8.63e+01</td></td></td>	4.7 <td>462</td> <td>1<td>US-08-471-</td><td>Sequence 20, Applicatio 8.63e+01</td></td>	462	1 <td>US-08-471-</td> <td>Sequence 20, Applicatio 8.63e+01</td>	US-08-471-	Sequence 20, Applicatio 8.63e+01
45	82 <td>4.7<td>462</td><td>2<td>US-08-469-</td><td>Sequence 20, Applicatio 8.63e+01</td></td></td>	4.7 <td>462</td> <td>2<td>US-08-469-</td><td>Sequence 20, Applicatio 8.63e+01</td></td>	462	2 <td>US-08-469-</td> <td>Sequence 20, Applicatio 8.63e+01</td>	US-08-469-	Sequence 20, Applicatio 8.63e+01

## ALIGNMENTS

RESULT	ID	1	STANDARD;	PRT;	403 AA.
XX	PCT-US95-05064-2	xxxxxx			
XX	Sequence 2, Application PC/TUS9505064				
DE	GENERAL INFORMATION:				
CC	APPLICANT: Corixa Corporation				
CC	TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE STIMULATION AND ENHANCEME				
CC	NUMBER OF SEQUENCES: 2				
CC	CORRESPONDENCE ADDRESS:				
CC	ADDRESSEE: SEED AND BERRY				
CC	STREET: 6300 Columbia Center, 701 Fifth Avenue				
CC	CITY: Seattle				
CC	STATE: Washington				
CC	COUNTRY: USA				
CC	ZIP: 98104-7092				
CC	COMPUTER READABLE FORM:				
CC	MEDIUM TYPE: Floppy disk				
CC	COMPUTER: IBM PC compatible				
CC	OPERATING SYSTEM: PC-DOS/MS-DOS				
CC	SOFTWARE: Patent Release #1.0, Version #1.30				
CC	CURRENT APPLICATION DATA:				
CC	APPLICATION NUMBER: PCT/US95/05064				
CC	FILING DATE: 24-APR-1995				
CC	CLASSIFICATION:				
CC	ATTORNEY/AGENT INFORMATION:				
CC	NAME: Kadlecsek, Ann T.				
CC	REGISTRATION NUMBER: P-39,244				
CC	REFERENCE/DOCKET NUMBER: 210121.404PC				
CC	TELECOMMUNICATION INFORMATION:				
CC	TELEPHONE: (206) 622-4900				
CC	TELEFAX: (206) 682-6031				
CC	TELEX: 3723836 SEEDANDBERY				
CC	INFORMATION FOR SEQ ID NO: 2:				
CC	SEQUENCE CHARACTERISTICS:				
CC	LENGTH: 403 amino acids				
CC	TYPE: amino acid				
CC	TOPOLOGY: linear				
CC	MOLECULE TYPE: protein				
CC	SEQUENCE 403 AA; 45269 MW; 838361 CN;				
CC	Query Match	5.5%;	Score 95;	DB 4;	Length 403;





[illegible]

APPLICATION NUMBER: US/08/607,509  
FILING DATE: 16-FEB-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MAKI, DAVID J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.404C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 403 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE 403 AA; 45326 MW; 829472 CN;

Query Match 5.5%; Score 95; DB 2; Length 403;  
Best Local Similarity 25.6%; Pred. No. 9,91e+00;  
Matches 21; Conservative 23; Mismatches 32; Indels 6; Gaps 6;

Db 230 ILVK-RESLTGIGIOEFLAVEEE-HKLTLDLYETVSIASVIRANRRKYDAENKI 287  
OY 66 LLKKRDITLMLND-KPFFVLIEDGTVEI-BEYFOALA-GDTV-PMYLQKGOKWOPPE 121

Db 288 NOSNHTVSMAHMEPKSDREY 309  
OY 122 OGTRHPDLSHKRPARKIDVARY 143

RESULT 6  
US-08-607-509-2 STANDARD; PRT; 403 AA.  
xxxxxx

Sequence 2, Application US/08607509  
Patent No. 5876735  
GENERAL INFORMATION:  
APPLICANT: Reed, Steven G.  
TITLE OF INVENTION: METHODS FOR ENHANCEMENT OF PROTECTIVE IMMUNE RESPONSES  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SEED AND BERRY LLP  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
City: Seattle  
STATE: Washington  
COUNTRY: USA  
ZIP: 98104-7092

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/607,509  
FILING DATE: 16-FEB-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MAKI, DAVID J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 210121.404C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 622-4900  
TELEFAX: (206) 682-6031  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 403 amino acids  
TYPE: amino acid





```
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1130 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: double  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: DNA (genomic)  
CC HYPOTHETICAL: NO  
CC ANTI-SENSE: NO  
CO SEQUENCE      1130 AA; 123655 MW; 6273599 CN;  
  
Db Query Match          5.0%; Score 87; DB 2; Length 1130;  
Best Local Similarity 28.0%; Pred. No. 3.81e+01;  
Matches   14; Conservative    13; Mismatches 22; Indels    1; Gaps    1;  
  
Oy        462 ATGAGDLEFSLGPPIYADVFVSHILKRPDRLLILDFAEEELADQFL 511  
           |||::||::| :: | : | : | : | : | : | : | : | : | :  
60 AASLEDLIRKV-RDTLMADKPFFLVLEEKGTVETEERFYALAGDVFM 108  
  
RESULT     13  
ID US-08-997-362-89 STANDARD; PRT; 376 AA.  
XX AC xxxxxx  
XX DT  
XX DE  
DE Sequence 89, Application US/08997362  
XX  
Sequence 89, Application US/08997362  
Patent No. 5985287  
GENERAL INFORMATION:  
APPLICANT: Tan, Paul  
APPLICANT: Hiyma, Jun  
APPLICANT: Visser, Elizabeth  
APPLICANT: Skinner, Margot  
APPLICANT: Scott, Linda  
APPLICANT: Prestidge, Ross  
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR  
TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS  
NUMBER OF SEQUENCES: 194  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Law Offices of Ann W. Speckman  
STREET: 2601 Elliott Avenue, Suite 4185  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98121  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/997,362  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970  
FILING DATE: June 12, 1997  
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347  
FILING DATE: August 29, 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Sleath, Janet  
REGISTRATION NUMBER: 37,007  
REFERENCE/DOCKET NUMBER: 11000.1002c2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-269-0565  
TELEFAX: 206-269-0563  
TELEX:  
INFORMATION FOR SEQ ID NO.: 89:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 376 amino acids  
TYPE: amino acid
```

[illegible]

Sun Jun 4 17:48:41 2000

US-09-471-749-3.rtf

Page 8

QY 82 FLVLEBGGTIVETEYEQALA-GD-TVFMVLQKQKQKQPPSEQTR-NPLSLSHKPAKKI 138

Db 74 NVNTTV-FOHVALPPH 87

QY 139 DVARVTFDLXKLNPO 153

Search completed: Sun Jun 4 15:28:28 2000  
Job time : 12 secs.

45 92 1070 2 US-08-633- Sequence 2, Applicatio 4.70e+01

ALIGNMENTS

RESULT 1 US-08-881-857-2 STANDARD: PRT: 405 AA.

Sequence 2, Applicatio US/08881857  
Patent No. 5919660  
GENERAL INFORMATION:  
APPLICANT: KIRKLY, KRISTINE K  
APPLICANT: GROSS, MITCHELL S  
APPLICANT: HURLE, MARK R  
TITLE OF INVENTION: HUMAN REQUITEM  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: RAINIER & PRESTIA  
STREET: P.O. BOX 980  
CITY: VALLEY FORGE  
STATE: PA  
COUNTRY: USA  
ZIP: 19482  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/881,857  
FILING DATE: 24-JUN-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/021,229  
FILING DATE: 26-JUN-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: PRESTIA, PAUL F  
REGISTRATION NUMBER: 23,031  
REFERENCE/DOCKET NUMBER: ATG-50013  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 610-407-0700  
TELEFAX: 610-407-0701  
TELEX: 846169  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 405 amino acids

Sun Jun 4 17:48:43 2000

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE 405 AA: 45780 MW: 803227 CN:  
14.7% Score 441: DB 2; Length 405;  
Best Local Similarity 41.0%; Pred. No. 7.37e-28;  
Matches 57: Conservative 30; Mismatches 48; Indels 4; Gaps 3;

DB 250 EGEDEKEDSOPPTPVSRSEEOKSKKGGPDGLALPNNYCDFCLGDSKINKKKTGQPEELVSC 309  
256 QKSKDKAAT-PRKDGPKKRSVLKSKSVFGYKPKVAPNAICIGICL-KGKESNKKKGKAKAESLIHC 313  
DB 310 SDGSGHSPSCLOFTPVMAAATKTYRMOCCINICGTCSENDOLLFCDDCDDRGYHMY 369  
314 SOCENSGHSPSCLOFTPVMAAATKTYRMOCCINICGTCSENDOLLFCDDCDDRGYHMY 373  
DB 370 CLTPSMSPPEEGSWSCNHC 388  
374 CV--GLGAIPIPSGRWICCC 390

RESULT 2 STANDARD: PRT: 405 AA.

US-08-881-857-2

\*\*\*\*\*  
M P E R E H  
(7M)  
\*\*\*\*\*

Release 3.1A John F. Collins, Bioinformatics Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.  
Distribution rights by Oxford Molecular Ltd

Mparch\_pp protein - protein database search, using Smith-Waterman algorithm  
R 1: Sun Jun 4 15:27:04 2000; Maspar time 17.96 Seconds  
625.036 Million cell updates/sec  
Tabular output not generated.

Title: >US-09-471-749-3  
Description: (1-238) from US09471749.ppe  
Perfect Score: 1742  
Sequence: 1 MEYAMKSLSLIPKSLSRHV.....GPPKGRKASLIPTCLKILQ 238

Scoring table:  
PAM 150  
Gap 11

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: plr62  
1:plr1 2:plr2 3:plr3 4:plr4

Statistics: Mean 45.595; Variance 83.289; scale 0.547

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Rd	Score	Query Match	Length	ID	Description	Pred. No.
1	1390	79.8	239	2 A42445	fat-specific protein	9.98e-281
2	102	5.9	1405	2 T04426	hypothetical protein	1.80e-01
3	101	5.8	335	2 H70107	hypothetical protein	3.40e-01
4	100	5.7	692	2 H70362	glycogen phosphorylas	6.35e-01
5	98	5.6	435	2 A70929	hypothetical protein	1.59e+00
6	97	5.6	919	2 B72765	hypothetical protein	1.59e+00
7	95	5.5	200	2 F70900	hypothetical protein	1.59e+00
8	95	5.5	403	2 T02861	EIF-4A protein - Leis	1.59e+00
9	95	5.5	413	2 S38358	translation initiatio	1.59e+00
10	95	5.5	1252	2 T14272	cortactin-binding pro	1.59e+00
11	94	5.4	341	2 C72403	hypothetical protein	3.92e+00
12	92	5.3	214	2 T08922	probable calcineurin	3.92e+00
13	92	5.3	254	2 S52021	translation initiatio	3.92e+00
14	92	5.3	413	2 S52018	translation initiatio	3.92e+00
15	92	5.3	414	2 JN0839	two-component sensor	3.92e+00
16	92	5.3	505	2 F69867	carboxylesterase (EC	3.92e+00
17	92	5.3	562	2 A55281	ATP-dependent permea	3.92e+00
18	92	5.3	633	2 S44795	ATP-dependent permea	3.92e+00
19	92	5.3	1049	1 S19421	TMY resistance protei	3.92e+00
20	92	5.3	1179	2 T04584	ABC transporter subu	7.04e+00
21	91	5.2	251	2 A69020	hypothetical 27k prot	5.26e+00
22	90	5.2	261	2 T00319	hypothetical protein	5.26e+00
23	91	5.2	298	2 T01948	hypothetical protein	5.26e+00

24	90	5.2	301	2 B75153	hypothetical protein	7.04e+00
25	91	5.2	316	2 D71535	probable NADH (ubiqui	5.26e+00
26	91	5.2	561	2 JC2447	carboxylesterase (EC	5.26e+00
27	90	5.2	674	2 JC5104	translation initiat	7.04e+00
28	91	5.2	2285	2 T12796	probable transglycosy	5.26e+00
29	88	5.1	261	2 T15520	hypothetical protein	1.25e+01
30	89	5.1	301	2 B69593	GTP-binding protein b	9.40e+00
31	88	5.1	405	2 I50608	bone morphogenetic pr	1.25e+01
32	88	5.1	412	2 JC1453	translation initiatio	1.25e+01
33	89	5.1	469	2 T07021	hypothetical protein	9.40e+00
34	89	5.1	587	2 T16617	hypothetical protein	9.40e+00
35	88	5.1	858	2 A46513	protein 4.1, p4.1-m	1.25e+01
36	88	5.1	870	2 F72058	DNA polymerase I - Ch	1.25e+01
37	89	5.1	896	2 S26740	gene borse protein - f	9.40e+00
38	89	5.1	896	2 A36455	bride of sevenless (b	9.40e+00
39	89	5.1	1082	2 T15269	hypothetical protein	9.40e+00
40	89	5.1	1102	2 S28104	probable DNA-directed	9.40e+00
41	88	5.1	1174	2 S28976	DNA-directed RNA poly	1.25e+01
42	88	5.1	1179	2 T05673	hypothetical protein	1.25e+01
43	88	5.1	1468	2 T05672	hypothetical protein	1.25e+01
44	87	5.0	348	2 C64676	oligopeptide ABC tran	1.66e+01
45	87	5.0	3951	1 VPIH81	Fl protein - avian in	1.66e+01

## ALIGNMENTS

RESULT 1  
ENTRY A42445 #type complete  
TITLE fat-specific protein FSP27 - mouse  
ORGANISM Mus musculus #common\_name house mouse  
DATE 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 11-Apr-1995

ACCESSIONS A42445  
REFERENCE A42445  
#authors Danesch, U.; Hoek, W.; Ringold, G.M.  
#journal J. Biol. Chem. (1992) 267:7185-7193  
#title Cloning and transcriptional regulation of a novel adipocyte-specific gene, FSP27. C/EBP-enhancer-binding protein (C/EBP) and C/EBP-like proteins interact with sequences required for differentiation-dependent expression.

#cross-references MIM:92202289  
#accession A42445  
#status preliminary  
#molecule\_type mRNA  
#residues 1-239 #label DAN  
##experimental\_source Fat cells  
##note sequence extracted from NCBI backbone (NCBIN:922164, NCBI:92165)

SUMMARY #length 239 #molecular\_weight 27292 #checksum 3042

Query Match 79.8% Score 1390; DB 2; Length 239;  
Best Local Similarity 79.1% Pred. No. 9.98e-281;  
Matches 189; Conservative 27; Mismatches 22; Indels 1; Gaps 1;

Db	1	MDYAMKSLILYPRSLSRHVA	STAVVTVQOLVSRSPSETPRARCSTADRYKRGIMA	60
Qy	1	MEYAMKSLILYPRSLSRHVS	RVSVTVQOLLSPSRKARAPCRVSTIDRSYKRGIMA	60
Db	61	HSLEDLTKVQDILKDKFSL	VEEDGTVETEEYFOALADTMFVLLKGKRRPS	120
Qy	61	YSLEDLTKRVDTLMDLDR	KFEVLDEGGTVETEEYFOALAGTVMVLOKQKRRPS	120
Db	121	EQRKRAQALBSQKPTKRI	IVARTFDLYKINRQDFGCLAVKATLBTSLSDLHCYK	180
Qy	121	EQRTRRP-LSLSRPARKI	IVARTFDLYKINRQDFGCLAVKATLBTSLSDLHCYK	179
Db	181	AKRIKELIVMTLFSKQAT	GHMLGTSSYMOQFLDADEEOPAPAKRSLLPCLIKMIO	239
Qy	180	AKRIKELAFKRALFPMQ	ATGHVILGTSCYQQLADAEEOQPRGRKASLIPTCLKILQ	238

RESULT 2



```
#molecule_type DNA  
#residues 1-335 #label KLE  
#cross-references GB:AEO0119; GB:AEO00783; NID:g2687936; PID:g2687947;  
TRIG.BB0063
```

##experimental\_source strain B31  
#length 335 #molecular\_weight 37305 #checksum 7981

```

Query Match          5.8%; Score 101; DB 2; Length 335;
Best Local Similarity 33.3%; Pred.No.2,48e-01;
Matches 17; Conservative 15; Mismatches 17; Indels 2; Gaps 2;

Db      26 TIKALLVIFGSLIVSFAIFFMVLNNETTPYPMNLSTLAIE-DAY-LDLOR 74
Oy      62 SLEDLLTKAVDTLMLADKRPFLVLEEDGTYYVEREEFYQLADGYFMVLQK 112
:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

RESULT      4
ENTRY       H70362 #type complete
TITLE      glycogen phosphorylase - Aquifex aeolicus
ORGANISM   #formal_name Aquifex aeolicus
DATE       08-May-1998 #sequence_revision 08-May-1998 #text_change
           08-May-1998

ACCESSIONS REFERENCE
#authors    H70362
            A70300
            Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, M.G.;
            Lenox, A.L.; Graham, D.E.; Overbeek, R.; Sned, M.A.;
            Keller, M.; Aubay, M.; Huber, R.; Feldman, R.A.; Short,
            J.M.; Olson, G.J.; Sanson, R.V.;
            Nature (1998) 392:353-358
            #journal
            #title The complete genome of the hyperthermophilic bacterium
                    Aquifex aeolicus.
            #accession H70362
            #cross-references MIMD:98196666
            #status preliminary; nucleic acid sequence not shown;
                    translation not shown

GENETICS
#gene       glpP
SUMMARY     #length 692 #molecular_weight 81158 #checksum 7682
            ##molecule_type DNA
            ##residues 1-692 ##label AOF
            ##cross-references GB:AEO00704; MID:g2983301; PID:g2983305; GB:AEO00657
            ##experimental_source strain VF5

Db      230 LRLKQQIYLGGHYTLLEKLGDIDAGGHINEDPSFVFIAEFILKKLGITWKAIEVR 289
Oy      68 LKVADTLMIA-DKRFELV-LEEDGTYYVEREYQNALAG-DYFMVLQKQKNQPPSEGT 124
:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

Db      290 KISLFTHHTPLR-YAANTYPPEM 311
Oy      125 RHPLSLSHKPAKKIDVARVTEDL 147
:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::
:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::

RESULT      5
ENTRY       A70929 #type complete
TITLE      hypothetical protein RV1783 - Mycobacterium tuberculosis
                (strain H37RV)
ORGANISM   #formal_name Mycobacterium tuberculosis
DATE       17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change
           17-Jul-1998

ACCESSIONS REFERENCE
#authors    A70929
            A70500
            Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher,
            C.; Harris, D.; Gordon, S.V.; Eigmeier, K.; Gas, S.; Barry,
            III, C.E.; Tekle, F.; Badcock, R.; Basham, D.; Brown, D.;
            Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.;
            Pettwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
            Hornsbury, I.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.;
            Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.;
            Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.;
            Skellern, S.; Squares, S.; Squares, R.; Sulston, J.E.;
            Taylor, K.; Whitehead, S.; Barrett, B.G.
            Nature (1998) 393:537-544
            #journal
            #title Deciphering the biology of Mycobacterium tuberculosis from
                    the complete genome sequence.
            #cross-references MIMD:982395987

```

```

#Journal
#title
#cross-references
#accession
#status
#molecule_type
#residues
#cross-references
#experimental_source
GENETICS
#gene
CLASSIFICATION
#superfamily
#length
SUMMARY
Query Match
Best Local Similarity
Matches
Db
51 WA-PVF-GIGAFILAHCFI
190 WAFSMQATSHVLGTCY
RESULT
ENTRY
TITLE
ORGANISM
DATE
ACCESSIONS
REFERENCE
#authors
#submission
#description
#accession
#status
#molecule_type
#residues
#cross-references
#accession
#status
#molecule_type
#residues
#cross-references
CLASSIFICATION
#superfamily
#length
SUMMARY
Query Match
Best Local Similarity
Matches
Db
230 ILVA-RESLLEGIKOFFIV

```

0Y 66 LLKRVDTLMAD-KPEFLVEEDGTVEET-EHYFOALA-GDTV-FMWLQKGOKWOPSE 121  
Db 288 NOSNTVSSMHAEMKRSDBRY 309  
0Y 122 OGTRHPLSLSHKPAKIDIVA 143

RESULT 9  
ENTRY S38358 #type complete  
TITLE translation initiation factor eIF-4A - rice  
ORGANISM #formal\_name Oryza sativa #common\_name rice  
DATE 19-May-1994 #sequence\_revision 10-Nov-1995 #text\_change 28-May-1999

ACCESSIONS S38358  
REFERENCE S38358  
#authors Nishii, R.; Kido, S.; Uchimiyu, H.; Kato, A.  
#journal Biochim. Biophys. Acta (1993) 1174:293-294  
#title Isolation and characterization of a rice cDNA which encodes the eukaryotic initiation factor 4A.

#cross-references MWID:93385159  
#accession S38358  
#status preliminary  
#molecule\_type mRNA  
#residues 1-413 #label NIS  
#cross-references GB:D12627; NID:g303843; PIDN:BA02152.1;  
PID:dl002642; PID:g303844

CLASSIFICATION #superfamily translation initiation factor eIF-4A; DEAD/H box  
#domain DEAD/H box helicase homology #label DEAD/  
#region nucleotide-binding motif A (P-loop) \\\  
#region nucleotide-binding motif B \\\  
#region DEAD motif  
#length 413 #molecular-weight 46931 #checksum 2148

SUMMARY  
Query Match  
Best Local Similarity 27.1%; Pred. No. 1.59e+00;  
Matches 19; Conservative 19; Mismatches 28; Indels 4; Gaps 4;

Db 242 ILVK-RDEFLTEIKQFYVNEKEKMLDPLDLYETLITQSVIFVTRKVMWLTDKMR 300  
0Y 66 LLKRVDTLMAD-KPEFLVEEDGTVEET-EHYFOALA-GDTVFWLQKGOKWOPSE 122  
Db 301 GDRHTVSATH 310  
0Y 123 GTRHPLSLSH 132

RE 10  
ENTRY T14272 #type complete  
TITLE cortactin-binding protein 1 - rat  
ORGANISM #formal\_name Rattus norvegicus #common\_name Norway rat  
DATE 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

ACCESSIONS T14272  
REFERENCE Z17952  
#authors Du, Y.; Weed, S.A.; Xiong, W.C.; Marshall, T.D.; Parsons, J.T.  
#journal Mol. Cell. Biol. (1998) 18:5838-5851  
#title Identification of a novel cortactin SH3 domain-binding protein and its localization to growth cones of cultured neurons.

#cross-references MWID:98414600  
#accession T14272  
#status preliminary; translated from GB/EMBL/DBJ  
#molecule\_type mRNA  
#residues 1-1252 #label DUY  
#cross-references EMBL:AF060116; NID:g3091151; PID:g3091152;  
PIDN:AC62226.1  
#length 1252 #molecular-weight 134280 #checksum 2350

SUMMARY  
Query Match 5.5%; Score 95; DB 2; Length 1252;

Best Local Similarity 28.1%; Pred. No. 1.59e+00;  
Matches 18; Conservative 18; Mismatches 26; Indels 2; Gaps 2;

Db 122 HLKRVTVTRNLDPDDTARKKAPPPKRAFTALTATVSKSTATLELYNDKKEEYVP 181  
0Y 19 HVSRTSVYVTOOL-LSESPKAPRARCVRSTADRSVRKGMVSELDLLKVR-DTLM 76

Db 182 ASKP 185  
0Y 77 ADKP 80

RESULT 11  
ENTRY C72403 #type complete  
TITLE hypothetical protein - Thermotoga maritima (strain MSB8)  
ORGANISM #formal\_name Thermotoga maritima  
DATE 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 11-Jun-1999

ACCESSIONS C72403  
REFERENCE A72200  
#authors Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Hatt, D.H.; Hickey, E.K.; Peterson, J.D.; Nelson, W.C.; Ketchum, K.A.; McDonald, L.; Utterback, T.R.; Malek, J.A.; Linher, K.D.; Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; Heidelberg, J.; Sutton, G.G.; Fleischmann, R.D.; White, O.; Salzberg, S.L.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
#journal Nature (1999) 399:323-329  
#title Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of Thermotoga maritima.

#cross-references MWID:99287316  
#accession C72403  
#status preliminary  
#molecule\_type DNA  
#residues 1-341 #label ARN  
#cross-references GB:AE001707; GB:AE000512; NID:g4980720; PID:g4980739;  
TIGR:TW0241

GENETICS  
#experimental\_source strain MSB8  
#gene TW0241  
#length 341 #molecular-weight 39568 #checksum 7991

SUMMARY  
Query Match  
Best Local Similarity 24.7%; Pred. No. 2.16e+00;  
Matches 18; Conservative 20; Mismatches 31; Indels 4; Gaps 4;

Db 138 LDYEDVSLVVKHEFELSTDEVEFKLQREDLRLNLEPVVLDPLN-GVT-IREIDVEER 195  
0Y 85 LREDDGTVEET-EHYFOALAGDTVFWLQKGOKWOPSEGTTRHPLSLSHKPAKIDIVA-R 142

Db 196 IDFVILNYGNEF 208  
0Y 143 VTFDLKLNPODF 155

RESULT 12  
ENTRY T08922 #type complete  
TITLE probable calcineurin B-like protein - Arabidopsis thaliana  
ORGANISM #formal\_name Arabidopsis thaliana #common\_name mouse-ear cress  
DATE 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 11-Jun-1999

ACCESSIONS T08922  
REFERENCE Z16518  
#authors Bevan, M.; Zimmermann, W.; Gruenelsen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Lemcke, K.; Schellier, C.  
#journal submitted to the Protein Sequence Database, May 1999

#accession T08922  
#status preliminary  
#molecule\_type DNA  
#residues 1-214 #label BEV



FEATURE

85-373 #domain DEAD/H box helicase homology #label DEAD\  
 85-92 #region nucleotide-binding motif A (P-loop)\  
 186-191 #region nucleotide-binding motif B\  
 190-193 #region DEAD motif\  
 91 #binding\_site ATP (Lys) #status predicted  
 SUMMARY #length 414 #molecular-weight 46927 #checksum 2762

Query Match 5.3%; Score 92; DB 2; Length 414;  
 Best Local Similarity 26.8%; Pred. No. 3.92e+00;

Matches 19; Conservative 20; Mismatches 27; Indels 5; Gaps 5;

Db 242 ILVK-RDELTLGSIKQFYVNEKEKIDPLCDLYETLAITQSIFVNTRRKYDMLTDKM 300  
 QY 66 LILKVDITMLAD-KPFFVLEEDGTVEFE-EYQALA-GDIV-FMYLQKQKQPPSE 121  
 Db 301 RGRDHTVSATN 311  
 QY 122 QGTRRPLSLSH 132

Se completed: Sun Jun 4 15:27:26 2000  
 Job time : 22 secs.